Industry Canada Industrie Canada

ECONOMIC RESEARCH AND POLICY ANALYSIS BRANCH

Working Paper Series

Venture Capital as a Catalyst for High Growth

Ryan Kelly, Industry Canada
Hankook Kim, Industry Canada

Working Paper 2013-01

Canada

Permission to Reproduce

Cat. No. Iu172-1/2013-1E-PDF

ISBN 978-1-100-22726-9

Venture Capital as a Catalyst for High Growth[†]

Ryan Kelly[*]

Economic Research and Policy Analysis Branch
Industry Canada
e-mail: ryan.kelly@ic.gc.ca

Hankook Kim

Economic Research and Policy Analysis Branch
Industry Canada
e-mail: hankook.kim@ic.gc.ca

September 2013

Abstract. We use new data linking information on venture capital (VC) financing from Thomson Reuters with firm-level administrative data from Statistics Canada to compare VC-backed and non-VC-backed firms. In particular, we assess the impact of VC on growth of key firm-level variables. This is the first study of its kind on VC-backed firms that incorporates official financial information of enterprises operating in Canada. The richness of the data allows us to construct a control group of non-VC-backed firms using previously unavailable covariates that are often thought to be associated with the potential for high growth and the probability of receiving VC financing, e.g., R&D expenditures, participation in an R&D support program, size, age, industry, location, among others. Using a propensity score matching estimator, our results suggest that VC-backed firms outperform their non-VC-backed counterparts in growth metrics. VC-backed firms experience higher growth in sales, wages, and employment over a 5-year period while the higher growth in R&D expenditures for VC-backed firms only occur over the first year with the levels converging over the long-term. In profitability, we find no statistical difference between VC-backed and non-VC-backed firms. Our results provide robust empirical evidence that VC financing is associated with faster firm growth and an acceleration of the innovation and commercialization process.

[†] The views and opinions expressed in this paper are those of the authors alone and do not represent, in any way, the views or opinions of the Department of Industry or of the Government of Canada.

[*] We would like to thank Statistics Canada, the Canadian Venture Capitalist Association and the Small Business Branch of Industry Canada for assistance and comments throughout this project. In particular, we would like to thank the efforts of Jim Valerio and Shane Dolan in spear-heading and managing the data development and descriptive analysis that made this research possible. We also owe special thanks to Anne-Marie Rollin and Younes Errounda for their tireless efforts in developing the final dataset. And finally, we would like to thank Thomas Hellmann and Leonard Sebatti for their comments.

1. Introduction

The significance of venture capital (VC) as a financing instrument for high-growth firms has been well documented. VC investors rely on specialized knowledge to identify early-stage firms with the potential for high growth, and proceed to buy large equity stakes in these firms with the objective of helping them to mature into profitable enterprises. As part of its growth strategy, VC provides not only financing but also close monitoring, expertise, and industry contacts to the investee firm. These additional supports could greatly improve a firm's chances for success given that VC-backed firms tend to be young and possess technologies, ideas, or assets with strong potential, but often lack the experience or the proven sales base to secure a sufficient level of financing to expand.

Given the managerial support provided through VC investments, it is not surprising to observe the impact of VC manifest itself through a number of different aspects of a firm's operation. In particular, the literature has identified VC's active role at the firm level in product market strategies or commercialization (Hellmann and Puri 2000), human resources practices (Hellmann and Puri 2002), and innovation strategies (da Rin and Penas 2007). These active roles have resulted in positive contributions measured at an aggregate level. Kortum and Lerner (2000) shows industries receiving more VC have greater patenting rates, and Samila and Sorensen (2011) identifies VC as a factor behind increased entrepreneurial activities in a metro area.

Anecdotally, the value proposition VC brings to the table—sales contacts, access to the new market, expertise in marketing technologies, among others—is fully consistent with the observation that VC funds could promote growth among the firms in which they invest. However, while the evidence in the literature demonstrates VC's different types of influences on a firm's operation, little is shown on a key aspect of VC investment, namely whether firms receiving VC grow faster than they would without this support.

To shed light on this issue, we use a newly constructed data set combining Thomson Reuters' VC data with Statistics Canada administrative data to examine the growth performance of firms backed by VC. Specifically, we construct a control group that mirrors our sample of VC-backed firms across a number of firm characteristics such as total assets, R&D expenditures, and age, among other characteristics that could affect firm growth outcomes. The covariates used in our model even include a dummy for participation in an R&D support program, i.e., Industrial Research Assistance Program (IRAP), which is likely to affect the probability of a firm seeking VC funding or focusing on growth.

Our results suggest that VC-backed firms enjoy a growth premium relative to non VC-backed firms following VC investment—VC-backed firms grow faster in terms of their assets, number of employees, revenue, sales and wages. We also find evidence that VC-backed firms increase their R&D expenditures more rapidly than comparable non VC-backed firms, but this result is limited to the near-term immediately following the first round of VC financing.

Our work is related to the literature assessing performance of VC-backed firms. An earlier example is Jain and Kini (1995) who find a positive and significant impact of VC monitoring on the post-IPO performance of VC-backed companies. Although the authors provide evidence that VC monitoring activities can improve firm performance, there are some limitations in interpreting their results. First, the study is limited to a sub-set of VC success stories reaching the IPO stage and rests on the assumption that VC maintains an active role in the firm post-IPO. Both of these concerns could overstate the importance of VC to firm performance. Further, by construction, the

study ignores firm performance prior to the IPO, which is arguably where VC support activities can have the most impact on firm performance.

A more recent effort, Puri and Zarutskie (2009), links a private source of VC data to official sources of firm financials in the United States from 1981-2001. They find that VC-backed firms grow faster relative to non VC-backed firms in terms of employment and sales, but are less profitable on a per-sales basis. However, given data constraints, the authors use a limited set of controls to select a control group for this comparison. While they include controls for important variables including size, age, location and industry, other potentially important variables describing the firm's finances and innovativeness are not accounted for in their methodology.

This paper adds to the literature in three ways. First, the richness of the data allows us to include a wide array of controls that could affect firm growth including measures related to the innovative performance of the firm, e.g., R&D expenditures. With these additional controls, we can construct a control group that reflects VC-backed firms more than in previous efforts. Second, this is the first empirical work of its kind using official financial information to assess the performance of VC-backed firms operating in Canada. And third, the timing of the data offers a more contemporary assessment of VC performance following an era of arguably excessive exuberance among investors in high-tech sectors.

The paper proceeds as follows: section 2 describes the data; section 3 explains our analytical framework and methodology; section 4 presents the results; and section 5 concludes with our findings.

2. Data

The data set was an ambitious undertaking that involved linking several external data sets to a number of administrative enterprise-level databases at Statistics Canada (STC Data). The administrative data sets include the Corporate Income Tax Returns (T2), Statement of Remuneration Paid (T4), and the Business Registry (BR).[1] These data were linked to external data from Thomson Reuters' VC database (Thomson Data) and the Industrial Research Assistance Program database (IRAP Data).

2.1 Data Linkage

Our sample is drawn from the population of incorporated enterprises employing at least one individual.[2] We refer to these units as firms in our analysis. Each database in the STC Data contains common enterprise-specific identifiers, which can be used to establish a deterministic linkage of all records across different databases in the STC Data. These data provide detailed enterprise-level information covering basic firm characteristics, typical items on financial statements, and employee information. Specifically, from the T2, we include Total Assets, Total Revenue, Sales, Net Income, Retained Earnings, Gross Profits, and R&D Expenditures; from the T4, we include the Total Payroll and the number of Individual Labour Units (ILUs);[3] and from

[1] The final data set also includes information from the PD7 payroll remittance slips, which offer more accurate employment measures. However, as the PD7 data is unavailable in 1999 and 2000, which covers approximately one third of the VC firm sample, we focus solely on the employment measures available in the T4 for our analysis.

[2] "The enterprise, as a statistical unit, is defined as the organisational unit of a business that directs and controls the allocation of resources relating to its domestic operations, and for which consolidated financial and balance sheet accounts are maintained from which international transactions, an international investment position and a consolidated financial position for the unit can be derived." http://www.statcan.gc.ca/concepts/definitions/ent-eng.htm

[3] The ILU is a continuous measure of the annual number of employees for a firm. Each individual (as identified by their Social Insurance Number) receiving a T4 represents 1.0 ILUs that are allocated to one or more firms according to the fraction of the

the BR we include the date of incorporation, the province of operation, and the primary NAICS code.

The STC Data are linked to two external data sets, Thomson Data and IRAP Data. The former provides information detailing the amount and date of all VC investments received by a given firm going back to the early 90's. This information is used to identify the treatment and control groups. The latter provides information on the amount and date of any IRAP funding a firm has received. Alongside the STC Data, this information helps us identify covariates for matching estimation. In particular, we anticipate IRAP-receiving firms are more likely to have a similar focus on technology and growth opportunities as those seeking VC investment. Further, given the similar support offered through IRAP funding, controlling for IRAP will reduce the possibility of attributing any difference in performance resulting from IRAP to VC.

The linkage of the STC Data to the external data was done based on names and addresses including postal codes and cities.[4] Not all firms in the BR had exactly the same names and addresses as their suspected counterparts in the Thomson Data and IRAP Data. Accordingly, some judgements were made to improve the linkage rate. For example, firms were considered the same if they have small variations in their names but identical addresses. Further, some of the linkages were established manually by experts familiar with the particular companies, i.e., analysts at the Small Business Branch of Industry Canada. After a rigorous linking exercise following a guiding principle of conservative accuracy, 1,044 VC-backed firms that received their first round of financing between 1999 and 2009 were identified in the STC Data. See Annex A for further information on the linkage between the STC Data and the external databases.

2.2 Further Refinement to Compare VC-backed to non VC-backed Firms

The linked data set contains longitudinal records for 1,044 VC-backed firms within our sample period. However, not all of these firms have suitable information for our analysis. In particular, several of the initial set of VC-backed firms receiving financing within the sample period did not file either a T2 or T4 in the year of their first financing round. Our approach involves matching VC-backed firms to those in the control group at the time of the first round of VC financing and examining their subsequent divergence in growth performance. As such, it is imperative to exclude firms without financial information at the time of first VC financing from our sample. In addition to this construct of the data, we impose an additional restriction that limits the treatment group to NAICS 4-digit industries where at least three firms have received funding at some point during the 1990-2009 sample provided by Thomson Reuters. This restriction is designed to mitigate any effects of firms with misclassified NAICs.[5] Following these adjustments to the data, we are left with a potential treatment group of 662 VC-backed firms; the sample size of the final matched treatment group will depend on the number of suitable analogues among the general population given the selected matching methodology.

---

individual's total annual payroll contributed by the particular firm. This measure will partially account for firms that hire employees later in the year as well as individuals working more than one job, but does not account for differences in the number of hours worked across employees.

[4] IRAP Data did not contain information on the operating street address or postal code, and as a result, was linked based on the enterprise name, and city / province. While the linkage was made using less information, the IRAP linkage achieved a higher initial rate of concordance due to higher quality matches among the firm names. This is not surprising as the recipients of IRAP funding are required to use their legal name and as a result, were more likely to be exactly identified in BR.

[5] Although ubiquitous for their ease of use, NAICS codes can be poor proxies for the market in which a firm competes as NAICS codes categorize firms based on production rather than on demand. VC investors focus on markets where they have technical knowledge or some form of expertise. Accordingly, a lack of investment activity among VC funds in a VC financed firm's NAICS code could indicate the assigned NAICS does not accurately reflect the firm's main market, or that the firm may not have suitable analogues within that NAICS code.

2.3 Longitudinalization

The initial linked data set is essentially a series of annual cross-sectional data over the 1999-2009 period, which is not adequate for our purpose as we intend to provide a robust assessment on growth performance of Canadian VC-backed firms. In the STC Data, firms are assigned a unique identifier in the BR (BRID), which is used to identify the firm in all databases. However, the assigned IDs can change over time for reasons that are unrelated to the destruction or creation of a firm, e.g., legal name change. Further, mergers and acquisitions are not identified or treated in the STC Data. To address these shortcomings and ensure growth performance is appropriately measured and attributed to the correct firm, the STC data are longitudinalized through a labour-tracking methodology. In particular, we track movements of employees based on T4 filings and corresponding employees' Social Insurance Numbers (SINs) to identify any structural change in a firm as an entity. For example, consider a case where we observe that a particular BRID associated with a firm that employs 50 people ceases to report any economic activity in one year. If a new BRID contemporaneously enters the sample that employs more than 25 of the individuals under the previous BRID, and we observe no other significant relationships with other BRIDs, we would conclude that the two BRIDs belong to the same firm, We use this procedure to create a single longitudinal record for each of the VC-backed firms as well as their matched counterparts in the control group using the year of first financing or year of the match as a base year. See Annex B for further information on the longitudinalization.

3. Matching VC-Backed and Non-VC-Backed Firms

We are interested in estimating the impact of VC financing and mentorship on various firm performance metrics. More formally, let $Y_i(1)$ represent the outcome, e.g., employment growth or R&D, for firm i receiving the treatment, i.e., having received VC financing, and let $Y_i(0)$ be the outcome of firm i without the treatment. Ideally, we would simply calculate the Average Treatment Effect on the Treated (ATT)[6] as shown in equation (1).

$$ATT = E[Y_i(1) - Y_i(0) \mid VC = 1] = \frac{1}{N_t} \sum_{i|VC=1} [Y_i(1) - Y_i(0)], \tag{1}$$

where N_t represents the number of treated firms, and VC is a dummy variable indicating whether a firm has received VC financing. However, as is generally the case with evaluation studies, we only observe one of these two states. Accordingly, the literature on measuring treatment effects approaches this problem as one of missing information. The standard practice is to account for the missing information by using data on non-treated firms in place of the unobservable counter-factual outcomes. In particular, one can use a general matching strategy to this end, and under certain conditions, it yields consistent estimates of the ATT.

A simple matching estimator requires two basic assumptions. First, there must exist a set of covariates, X, such that firm outcomes are independent of the treatment after controlling for these characteristics. Given that we are specifically interested in estimating the ATT, we can weaken this assumption to mean independence. See Heckman, Ichimura and Todd (1998) for further information. Second, we assume for all X, there is a positive probability of either receiving or not receiving VC financing. In effect, the second assumption stipulates there exist suitable analogues

---

[6] Given only a small sub-sample of the Canadian firm population is considered a viable target for VC financing, we do not estimate the Average Treatment Effect (ATE) in the analysis as it produces an average effect over the entire firm population, which is of little interest and difficult to interpret meaningfully.

for the treated firms within the general population to create an adequate control group. More formally, these assumptions can be written as:

Assumption 1: Unconfoundedness or Ignorability - $E[Y(0), Y(1)] \perp (VC \mid X)$

Assumption 2: Region of Common Support - $0 < \Pr(VC = 1 \mid X) < 1$

If these assumptions are satisfied, equation (2) yields consistent estimates of the ATT.

$$ATT = \frac{1}{N}\sum_{i=1}^{N}[(Y_i(1) \mid VC = 1, X = x) - (Y_i(0) \mid VC = 0, X = x)], \qquad (2)$$

where the second term in the summation operator describes the mean outcome of a control group matched to the sample of VC-backed firms using the set of covariates, X. Starting from assumptions 1 and 2, one can implement several different matching methods available to construct the control group on the set of covariates.

We first test to determine how our sample of VC-backed firms at the time of their first financing compares to the general population of firms in Canada over the 1999-2009 sample period. A straight comparison between averages for these respective groups will not be very informative. In contrast to the general population, the majority of the firms in the VC-backed sample are concentrated near the beginning of the sample period. Further, the VC-backed sample exhibits a different industry distribution which is disproportionately concentrated within professional services. To address these concerns, we calculate the average ratio of the values for the VC-backed firms relative to the average value for their corresponding 4-digit NAICS category and year of financing. In Table 1, we report these ratios and whether they are statistically different from one. The results show our sample of 662 VC-backed firms is significantly different from the general population of firms in Canada across several covariates that could be important in determining firm growth outcomes. Accordingly, matching methods are necessary to construct an appropriate control group.

Table 1: Benchmarking VC-Backed Firms to General Population During the First Year of VC Financing

Variable	Times Larger than General Population*	# of observations	p-value
Total Assets	2.36	662	0.000
Sales	1.38	662	0.002
Employment	2.43	662	0.000
Wages	1.16	662	0.000
Age	0.61	662	0.000
Gross Profits	1.59	662	0.000
Gross Margin**	0.77	482	0.000
Gross Profits per Employment**	0.59	485	0.000
R&D Expenditures***	1.82	432	0.000

* mean { $X(VC_{f,i,t})$ / mean { $X(\text{non-}VC_{f,i,t})$ } }

f-firm, i-NAICS 4 digit, t-year

The General Population refers to all enterprises operating in Canada reporting the necessary financial information and operating in the relevant 4-digit NAICS codes during the 1999-2009 sample (2,573,663 observations).

** Means based on firms with positive profits

*** Means based on R&D performers only

3.1 Exact Matching

The simplest and most data-intensive matching method is *exact matching*. Specifically, this strategy involves pairing treatment observations with exact matches among all desired covariates from the pool of non-treated firms. While *exact matching* provides the most robust matching results, its implementation, particularly over a large set of covariates or continuous covariates, is highly data-intensive and often not a viable option when using real-life data. Researchers could choose to reduce the set of covariates or categorize them to implement a variant of *exact matching*; however, this may increase the likelihood of violating the ignorability assumption.

In the literature on the performance of VC-backed firms, Puri and Zarutskie (2009) use *exact matching* as part of their empirical strategy. However in doing so, they match firms based on a small number of mostly discrete covariates, i.e., age, industry classification, geographical region, and employment size. In this paper, we perform an *exact matching* exercise to illustrate our approach in relation to Puri and Zarutskie (2009); however, our results suggest that *exact matching* may not be the optimal approach to explore our data. In particular, given our research interest, we opt to match firms based on financial variables at their initial level and examine the growth of these variables following investment. This leads to matching based on a large set of continuous covariates, and there is simply not enough data to include all the pertinent firm characteristics while achieving exact matches. Our results particularly highlight how matched firms can be nearly identical in a small set of measures, yet significantly different in other related and potentially important variables. This is further discussed in section 4, where we present empirical results.

3.2 Propensity Score Matching

Given the objective of our research and the richness of the data available, a matching strategy better suited for a large set of continuous covariates than *exact matching* is required. We use propensity score matching as first outlined in Rosenbaum and Rubin (1983). This strategy involves estimating the probability of receiving treatment (propensity score) and matching

treatment observation to potential controls with a similar propensity score. In effect, the propensity score summarizes the covariate distribution for each observation while accounting for each variable's significance in explaining the treatment assignment. It follows that matching on the propensity score allows us to include more firm characteristics as we can create a control group with a similar distribution among the desired covariates without requiring exact (or near exact) values for each variable.

Our propensity score matching estimator involves three steps.

First, we fit a logit model to calculate the predicted probability of receiving VC for the entire population, as in equation (3).

$$\Pr(VC_{it} = 1) = \alpha_c + \alpha_t + \alpha_{ind} + \beta X_{it} + \varepsilon_{it}, \tag{3}$$

where X_{it} is the full set of the desired covariates, α_c is the constant, α_t and α_{ind} are time- and industry-specific intercepts, and ε_{it} is the standard econometric error term.

Second, we define tolerance thresholds or 'calipers' using the propensity score calculated in the first step to create a set of potential matches for each treatment observation that avoids poor matches. Although the selection of this threshold is somewhat *ad hoc* by design, there is some guidance available for practitioners. In particular, Rosenbaum and Rubin (1985) suggest a general starting point of 0.25 standard deviations of the linear propensity score, which we adopt for this matching exercise as described in equation (4). If there is no match for an observation from the group of VC-backed firms that satisfies the predefined threshold, it is dropped from the sample as it has no suitable analogue among non-VC-backed firms.

$$\delta = 0.25 * \sigma \left\{ \ln \left(\frac{\Pr(VC = 1)}{\left(1 + \Pr(VC = 1)\right)} \right) \right\}. \tag{4}$$

Third, we construct the control group by pairing each treatment observation to the corresponding observation within the set of potential matches described in the second step that is closest in terms of the linear propensity score. More explicitly, let p describe the linear propensity score and I_0 the set of non-VC-backed firms within the same 4-digit NAICS code and province during the first year of VC financing. The match for firm i, M_i, is defined as:

$$M_i = \left\{ \min_j \left\| abs(p_i - p_j) \right\|, j \in I_0 \mid abs(p_i - p_j) < \delta \right\}. \tag{5}$$

We remove each pair of VC-backed and non-VC-backed firms from the pool once they are matched. In matching terminology, we follow a nearest-neighbour matching strategy on the linear propensity score without replacement. Given that some of the non-VC-backed firms could be the closest match for multiple VC-backed firms, the order in which the matches are made could impact which firms are selected for the control group. As several of the treatment observations had a large pool of potential matches, optimal matching procedures that minimize a global distance measure are too computationally onerous and beyond the scope of this paper.

Accordingly, we order the treatment firms for matching based on their date of first financing to maximize the potential length of the longitudinal records for the treatment and control groups.[7]

It is important to note that the preceding methodology can only account for the observable differences between the treatment and control groups as indicated by the ignorability assumption. However, as noted in Lerner (2010), the average VC investor will often take over 100 hours to screen a potential investment. Attempting to mitigate all the differences between the sample of VC-backed and non-VC-backed firms using purely administrative data is likely an impossible endeavour regardless of the selected methodology. However, while we can not completely separate firm outcomes resulting from selection issues from the impact of VC funding and mentoring, given the scope of our matching as well as the new types of covariates used in this paper, our results provide valuable economic perspectives on the growth performance of VC-backed firms.

4. Results

4.1 Exact Matching

We implement an *exact matching* estimator similar to Puri and Zarutskie (2009). We match on four covariates: industry as in 4-digit NAICs, location as in province of operation, age as in years since incorporation, and size as in number of ILUs from the T4. Puri and Zarutskie (2009) use slightly different industry and location classifications. Further, they use annual headcounts to measure firm size while we use a more refined employment measure that partially captures part-time workers or those joining the firm mid-year. Since our ILU measure is continuous and contains decimal figures, we use a 10 per cent rule where firms are flagged as a potential match if their ILU is either within 10 per cent of the number of ILUs of the treatment firm or within 1 ILU if the treatment firm has less than 10 ILUs. This matching exercise results in 377 matched pairs or a 57 per cent match rate.

Table 2 compares the treatment and control groups over several covariates. Since *exact matching* by definition will result in exactly the same location and industry, there is no need to compare the sample over those covariates. The upper section of the table compares the pairs over other covariates used in the matching and as expected, the matched pairs share very similar employment levels and an identical age profile. The bottom section of the table shows the comparison over other variables of interest that could affect VC financing decisions and/or firm growth outcomes. While *exact matching* produces very close matches over a small number of covariates, the results in the bottom section of the table show that the matched sample of VC-backed firms is significantly different from the sample of non VC-backed firms in all the other measures. The differences are dramatic in a statistical sense as all the p-values of the differences are zero at the three-decimal level, which provides strong statistical evidence that the matched control group differs from the treatment group over these measures.

The results raise significant concerns over whether *exact matching* based on these covariates is the best empirical approach for this paper. Some of the variables over which there is a significant difference between a paired match could be critical in assessing growth performance of VC-backed firms. On average, the VC group has more capital and employs higher skilled workers (as proxied by wages), but shows lower revenues and sales than the control group. Given these initial differences, subsequent observations on growth could be difficult to interpret. The initial lower

[7] Although not reported, matching with replacement yielded similar results suggesting low potential gains from optimal matching methods in this context.

level of sales among VC-backed firms, despite higher capital and labour costs, could for example, suggest that the firms in the control group tend to have a relatively more mature sales base and accordingly may be less likely to experience high growth in the near-term.

Table 2: Comparison between Matched Pairs - Exact Matching

Covariate	Mean VC	Mean Control Group	Difference	P-value of Difference	Standardized Difference in Means
		Covariates Used			
Ln Employment	2.193	2.185	0.008	0.931	0.007
Age	3.694	3.694	0.000	1.000	0.000
		Other Pertinent Measures			
Ln Total Assets	14.081	12.949	1.132	0.000	0.749
Asinh Sales	9.877	13.141	-3.264	0.000	-0.531
Ln Wages	10.754	10.426	0.328	0.000	0.620
Asinh Retained Earnings	-10.825	0.704	-11.529	0.000	-1.247
Asinh Revenue	12.389	13.953	-1.563	0.000	-0.338
Asinh Net Income	-9.926	1.156	-11.082	0.000	-1.159
Asinh R&D Expenditures	9.435	3.212	6.223	0.000	1.034
Asinh Gross Profits	8.346	11.851	-3.506	0.000	-0.457

Standardized difference in means = (mean_X[VC] - mean_X[Control]) / Std. Dev_X[VC]
Asinh refers to the inverse hyperbolic sine transformation.
VC obs = 377; Control group obs = 377

In Table 3 we focus specifically on the balance between the treatment and control groups in terms of counts of firms showing positive values for two indicators of innovative activity. R&D expenditures are an input to the innovation process and are often associated with firm growth, innovation and productivity growth. The exact matching procedure results in a poor balance in both the average R&D expenditure (shown in Table 2) and the number of R&D performers shown in Table 3.

Another potential indicator for firm strategies orientated towards innovation is whether the firm has received IRAP funding. To receive IRAP funding, firms must pass a vetting process whereby they must demonstrate that their proposed project or business plan is commercially viable and sufficiently innovative. Beyond the technical aspects of the proposed project, the due diligence process for IRAP funding also considers the business and management capabilities of the firm. Although the firm support offered by IRAP differs from VC—IRAP supplies direct funding as opposed to taking up equity stakes—it offers similar mentoring services and targets similar types of firms. Accordingly, whether a firm receives IRAP funding is arguably one of the most robust characteristics to assess the similarity between a VC-backed firm and a potential control. However, as in the case of R&D expenditures, the control group resulting from exact matching fails to achieve a similar balance in the number of IRAP recipients. The VC treatment group has nearly four times the number of IRAP recipients as the control group, suggesting the VC group is more focussed on the development and commercialization of innovative products. These results further question whether the growth prospects of these two groups are similar at the time of the match.

Table 3: Comparison between Matched Pairs - Exact Matching for R&D and IRAP

	Count		Percentage	
	VC	Control Group	VC	Control Group
R&D performer in year of match	242	87	64	23
R&D performer during sample period	277	123	73	33
IRAP recipient in year of match	76	18	20	5
IRAP recipient during sample period	116	23	31	6

VC obs = 377; Control group obs = 377

To make robust comparisons, the treatment and control groups need to be similar across all covariates that could impact performance outcomes. In contrast, our *exact matching* methodology produced matched treatment and control groups with significant imbalances across several of these covariates. This issue cannot be solved by simply extending the number of covariates for the matching process, as it would be too data intensive to achieve exact matches across all the detailed firm characteristics at our disposal while maintaining a reasonable match rate. Accordingly, we do not pursue this methodology further as the evidence provides little confidence that the control group satisfies the ignorability assumption, which if violated, results in inconsistent estimates of the average treatment effect on the treated.

4.2 Propensity Score Matching

Using propensity score matching first requires us to estimate a model to produce predicted probabilities of receiving treatment. We report the results from the logit estimation of propensity score in Table C1 of Annex C. To increase the likelihood that the ignorability assumption is satisfied, we include all available variables that could affect either firm outcomes or the probability of receiving VC financing. We also include the corresponding squared values for all the continuous variables to allow for non-linear relationships. The variables used in the logit estimation are total assets, sales, number of ILUs, wages, retained earnings, revenues, net income, age, R&D expenditures, a dummy variable indicating whether the firm received IRAP funding, a dummy variable for firms financed in 1999 that conducted R&D in 2000,[8] industry fixed effects based on 4-digit NAICs, and year fixed effects.

Our matching strategy based on propensity scores leads to an 82 per cent match rate, resulting in 544 matched pairs. Table 4 shows the differences between the matched samples of VC-backed and non-VC-backed firms. They are very similar in all dimensions and the p-values of the differences are high enough to suggest the two groups of firms are not statistically different. In addition to standard hypothesis testing, we present the results from an alternative balance measure—the standardized difference in means (SDM). This measure is often considered more reliable in the context of assessing changes in balance between two samples as standard tests are subject to a higher probability of encountering type II errors when the sample size is reduced. Rubin (2001) suggests that SDM values exceeding 0.5 can become problematic. As shown in Table 4, the SDM values for all of our covariates of interest and that of the linear propensity score are at or below 0.05, suggesting the treatment and control groups are well balanced.

[8] R&D expenditure data is not available for the year 1999. Note that we do not include firm observations from 1999 when calculating growth rates for R&D in the subsequent section. The dummy variable is for the sole purpose of achieving a better match among innovative firms financed in 1999.

Table 4: Comparison between Matched Pairs—Propensity Score Matching

Covariate	Mean VC	Mean Control group	Difference	P-value of Difference	Standardized difference in means
Ln Total Assets	14.485	14.453	0.032	0.735	0.020
Asinh Sales	11.618	11.835	-0.217	0.539	-0.037
Ln Employment	2.717	2.765	-0.049	0.544	-0.036
Ln Wages	10.683	10.674	0.009	0.787	0.016
Asinh Retained Earnings	-8.089	-7.835	-0.254	0.727	-0.021
Asinh Revenue	13.507	13.454	0.053	0.850	0.012
Asinh Net Income	-7.572	-7.305	-0.267	0.707	-0.023
Age	4.952	5.042	-0.090	0.796	-0.016
Asinh R&D Expenditures	8.228	8.456	-0.227	0.566	-0.035
Asinh Gross Profits	9.617	9.201	0.416	0.410	0.052
Linear Propensity Score	-4.729	-4.754	0.025	0.860	0.011

Standardized difference in means = (mean_X[VC] - mean_X[Control]) / Std. Dev_X[VC]

Asinh refers to the inverse hyperbolic sine transformation.

VC obs = 544; Control group obs = 544.

The degree to which the covariate balance is improved through propensity score matching is further highlighted through comparing the pairs in terms of the number of R&D performers and IRAP recipients. Table 5 shows these results, which reinforce our confidence in the control group. Both the treatment and control groups have nearly identical counts of R&D performers and IRAP recipients, suggesting a similar degree of focus on innovative strategies across the treatment and control groups.

Table 5: Comparison between Matched Pairs—Propensity Score Matching for R&D and IRAP

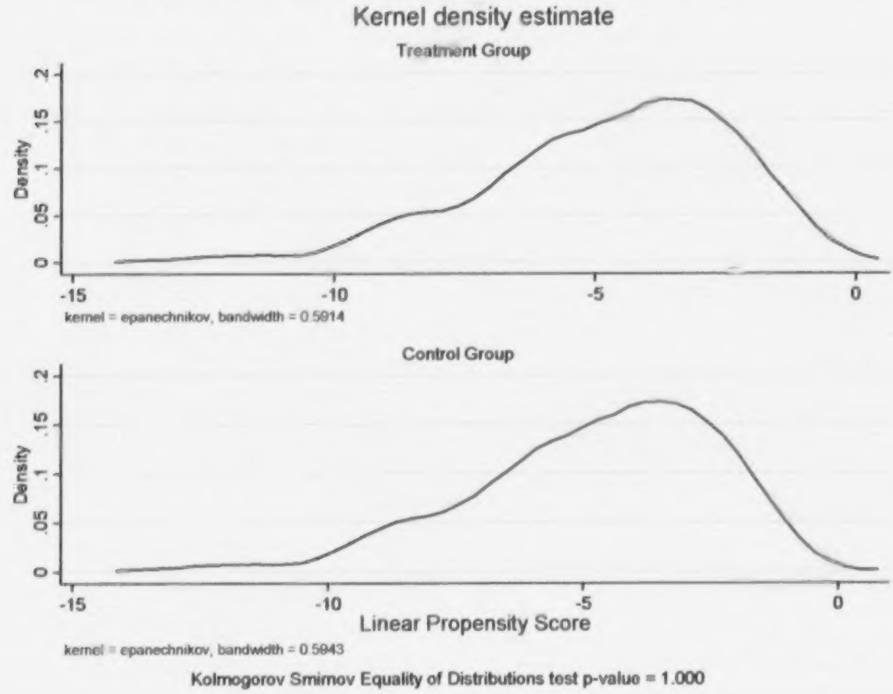
	Count		Percentage	
	VC	Control Group	VC	Control Group
R&D performer in year of match	336	347	62	64
R&D performer during sample period	412	396	76	73
IRAP recipient in year of match	91	88	17	16
IRAP recipient during sample period	152	156	28	29

VC obs = 544; Control group obs = 544.

Moving beyond the mean comparisons and simple firm counts, we compare the actual distribution of the linear propensity score for the treatment and control groups in Figure 1.[9] The kernel density estimates show that the treatment and control groups have a nearly identical distribution for the linear propensity score. This visual inspection is confirmed by the Kolmogorov Smirnov equality of distributions test, which has a p-value very close to one, suggesting that the two distributions are indistinguishable in statistical terms.

[9] Unfortunately, we cannot replicate this exercise for other variables while maintaining Statistics Canada's confidentiality requirements.

Figure 1: Kernel Density Plot of the Linear Propensity Score



The results in Tables 4, 5 and Figure 1 show our matching estimator is robust as both samples share similar characteristics over important dimensions such as their investment in R&D, size, age, and profitability, among other characteristics. Unlike our initial matched sample constructed using *exact matching*, the firms are broadly similar over all the covariates, and particularly among those that are pertinent to our research questions such as size and tendency to do R&D.

4.3 Growth Comparisons

With a well-matched sample of VC-backed and non-VC-backed firms, we can compare their respective growth rates over various timeframes. We compute the desired growth rates as follows:

$$y\ year\ Growth = \frac{\sum_{i=1}^{n}\left(\ln X_{i,t+y} - \ln X_{i,t}\right)}{n}, \qquad (6)$$

where y indicates the number of years span used to compute the growth rate. We compute growth rates over one-, three-, and five-year intervals. Commensurate with calculating growth, we only use firm observations that have positive values in both periods under consideration. In addition to this construct, we ignore firm values if the firm experienced an exit (i.e., merger or acquisition) prior to the end of the period. See Annex B for more details on the identification of these exits.

All the performance measures referring to dollar amounts are in nominal terms as industry and time specific price deflators are unavailable in our data. As a result, all growth rates, except for those relating to employment or the gross margin, reflect nominal growth. Despite this limitation, the growth comparisons between VC-backed and non VC-backed firms are still valid as, by

construction, the treatment and control group have identical year and industry profiles at the time of their match, and thus should have faced similar price inflation over the sample period.

However, there are two potential complications. First, given we are measuring growth rates up to five years, attrition in the sample could disrupt this initial year and industry balance if there are asymmetrical exits among the matched pairs of treatment and control firms. And second, the growth rate comparisons could be influenced by a survivor bias if, for example, poorly performing firms are more likely to remain in the sample among either the control group relative to the treatment group or *vice versa*. To address these issues, we replicate the results limiting the sample to firms where both components of the matched pair report the necessary information needed to calculate their growth rates. The results of this exercise yielded results that were qualitatively and quantitatively similar to our main results, albeit with slightly less statistical significance among a minority of the growth comparisons. See Table D1 of Annex D for these results.

In keeping with our main research question, we compare the growth rates between the treatment and control groups for four measures of firm scale: employment, total assets, revenue and sales. As shown in Table 6, the VC treatment group experienced substantial growth after the match. The highest annual growth rates occur in the first year following investment; the one year growth rate varies by indicator and ranges from approximately 20 per cent growth in total assets to 50 per cent growth in total revenue. This initial growth in scale seems to level off after the first year for some measures. For example, the cumulative growth rates in employment for VC-backed firms over 3 and 5 years are 41.5 per cent and 50.6 per cent, which are a slight increase from the 1-year growth rate of 34.4 per cent. While the employment figures can be interpreted as real growth, other growth rates are in nominal terms and include inflation. Accordingly, one needs to assess the difference in growth rates between VC-backed and non-VC-backed firms to assess the role of VC financing in promoting growth.

Table 6: Growth Rates for the Treatment and Control Group, Firm Scale Measures

	Mean VC	Mean Control Group	Difference	P-value of Difference	# VC	# Non-VC
Total Assets						
1 Yr growth	19.6%	0.1%	**19.5%**	0.001	487	493
3 Yr growth	38.7%	-10.2%	**48.8%**	0.000	334	360
5 Yr growth	53.5%	1.1%	**52.4%**	0.001	215	254
Employment						
1 Yr growth	34.4%	6.6%	**27.7%**	0.000	496	485
3 Yr growth	41.5%	-3.5%	**45.0%**	0.000	322	337
5 Yr growth	50.6%	3.6%	**47.0%**	0.000	193	225
Total Revenue						
1 Yr growth	52.9%	28.6%	**24.2%**	0.002	470	468
3 Yr growth	93.7%	36.7%	**57.0%**	0.000	323	333
5 Yr growth	107.3%	40.3%	**67.0%**	0.001	203	232
Sales						
1 Yr growth	48.7%	27.2%	**21.6%**	0.022	383	403
3 Yr growth	100.1%	47.2%	**53.0%**	0.001	262	273
5 Yr growth	137.4%	56.0%	**81.4%**	0.001	156	181

In comparing growth rates between the control and treatment groups, our results suggest that VC-backed firms grow faster than non VC-backed firms regardless of which measure is used, or the time horizon that we select to measure growth. Among all the performance comparisons, the growth premiums for VC-backed firms relative to non VC-backed firms were statistically significant at the 0.01 level with the exception of one value that was significant at the 0.05 level. The magnitude of the estimated growth premium for VC-backed firms is roughly similar across measures. Among the five-year cumulative growth rates, the estimated growth premium ranges between 47 per cent and 81.4 per cent, with the highest premium occurring for sales. The consistency across these growth measures strengthens the conclusion that VC-backed firms grow faster than non VC-backed firms. The statistically significant differences between these growth rates suggest the observed association between VC financing and growth is not inflation-driven. Further, the revenue and sales results suggest that VC-backed firms do not simply just purchase assets or hire employees following VC investment—they often successfully grow their business. On average, surviving VC-backed firms increase their sales, whether through expansion to meet demand in their current market or by successfully entering new markets to tap additional sources of sales.

Firms looking to develop and commercialize new goods and services are a prime target of VC investors. These innovative activities are likely a focal point of VC monitoring and mentorship support. Accordingly, we also investigate growth in two performance measures related to these processes: wage rates and R&D expenditures.

Starting with wages, the development and commercialization stages often require highly specialized labour. Meeting these requirements for high value-added employment to successfully commercialize new products should raise the average wage rate for the firm. However, this effect could be dominated by other changes within the firm's labour profile, e.g., the large overall increases in employment related to strategies geared towards firm expansion observed in Table 6.

We report the wage rate growth comparison between the treatment and control group in Table 7. The VC-backed firms had a 3.6 per cent wage growth premium over non VC-backed firms in the year following financing, and a 10.1 per cent cumulative growth premium over the 5-year period. The wage results are statistically significant for the 5-year growth at the 0.05 level while the 1-year growth rate is statistically significant at the 0.1 level. The 3-year growth rate is not quite statistically significant at the 0.1 level. The statistically significant difference in wage growth over the long term, i.e., five years, suggests wage growth is not purely driven by the injection of new funds in the form of VC investment and that real value creation through the utilization of the labour seems to take place after VC has invested.

Table 7: Growth Rates for the Treatment and Control Group, Wages and R&D Expenditures

	Mean VC	Mean Control Group	Difference	P-value of Difference	# VC	# Non-VC
Wages						
1 Yr growth	8.2%	4.6%	**3.6%**	0.061	496	485
3 Yr growth	16.7%	11.8%	4.9%	0.106	322	337
5 Yr growth	29.4%	19.3%	**10.1%**	0.029	193	225
R&D Expenditures						
1 Yr growth	25.3%	9.2%	**16.1%**	0.012	292	294
3 Yr growth	24.9%	8.9%	16.0%	0.230	178	172
5 Yr growth	48.9%	29.4%	19.5%	0.230	99	103

In growth of R&D expenditures, the VC-backed firm sample initially shows a high growth premium that is statistically significant; approximately 16 per cent in the first year. However, the observed growth premium for VC-backed firms is not statistically significant over the three- and five-year periods. This lack of statistical significance over longer timeframes could reflect a higher importance of R&D in the initial development and commercialization stage to meet firm growth objectives. In particular, given the previous result showing a sales growth premium among VC-backed firms, our results on R&D expenditures suggest VC involvement could accelerate the commercialization process, but may not necessarily lead to more innovation over the long term. This result is consistent with Hellmann and Puri (2000) who find that VC investment is associated with a faster time to market; and with Engel and Keilbach (2007) who find that after controlling for the number of patent applications at firm foundation, VC does not significantly affect firms' patenting behaviour. It is important to note that this interpretation of our results does not necessarily contradict Kortum and Lerner (2000), who find a positive impact of VC funding on patenting rates. Their results are at the industry level, and the impact of VC funding on industry patenting rates could work through other channels, such as an increase in entrepreneurial activity as found in Samila and Sorensen (2011).

Lastly, in Table 8 we compare the profitability of VC-backed and non-VC-backed firms using two measures: gross margin and gross profits per employee. We find no statistically significant divergence in profitability between the treatment and control groups over any period of time. The result that gross profits per employee are statistically similar suggests, at least for surviving VC-backed firms, the rapid increase in size has had little negative impact on their capacity to generate profits over the long term. This result sheds further light on the impact of VC on profitability and suggests a neutral effect of VC on firm profitability, which differs slightly from the findings of Puri and Zarutskie (2009). In particular, using the U.S. data, they find that VC-backed firms grow faster in terms of employment and sales but that profitability among VC-backed firms drops following VC investment. That said, our results do not provide robust evidence on the positive impacts of VC on profitability, and as such, they are consistent with the prevailing thought in the literature that VC's impact is more pronounced on growth than profitability.

It is important to note that these results do not construe as evidence regarding the performance or return of VC funds in Canada. The rate of return for a VC's investment in the firm depends on the exit value of the firm; and given that VC targets small early-stage firms, the exit value for a VC investment is likely more related to growth in firm scale rather than changes in profits relative to sales or employment.

Table 8: Growth Rates for the Treatment and Control Group, Profitability Measures

	Mean VC	Mean Control Group	Difference	P-value of Difference	# VC	# Non-VC
Gross margin						
1 Yr growth	-1.2%	-0.5%	-0.7%	0.717	341	338
3 Yr growth	0.5%	6.4%	-5.9%	0.211	240	243
5 Yr growth	0.8%	5.7%	-5.0%	0.255	138	159
Gross Profit/Employment						
1 Yr growth	14.3%	26.7%	-12.5%	0.183	341	333
3 Yr growth	54.6%	48.1%	6.4%	0.668	235	233
5 Yr growth	70.2%	71.3%	-1.1%	0.958	135	152

5. Conclusion

In this study, we compare VC-backed and non VC-backed firms across several firm performance measures using a unique dataset allowing us to control for firm characteristics unavailable in previous empirical works. In addition to a wide array of pertinent financial and employee information, we control for innovative measures including R&D expenditures and direct government support for innovation and commercialization. While we cannot completely eliminate selection effects of VC selecting firms with good potential—nothing can—the breadth of firm characteristics used for controls and magnitude of the estimates suggest that our results provide valuable economic perspectives regarding the potential impact of VC on growth.

In particular, our results suggest that VC support activities focus on expanding the size of the investee firm. We compare VC-backed and non VC-backed firms across four measures of firm

scale: total assets, employment, revenue and sales. Regardless of the measure, VC-backed firms enjoy a statistically significant growth premium over non-VC-backed firms. We also find evidence suggesting VC-backed firms have higher average wage growth rates, an indication that they add more high value-added employment than non-VC-backed firms. In terms of R&D expenditures, our results show VC-backed firms increase their expenditures relative to similar non VC-backed firms; however, the results are statistically significant only for the first year immediately following their first round of VC financing.

Finally, our results suggest VC-backed firms do not outperform their non VC-backed counterparts in terms of profitability. This observation is consistent with the prevailing thought on the industry and the literature that VC focuses on growth over profitability.

On balance, our results provide robust empirical evidence that VC financing is associated with both faster firm growth and an acceleration of the innovation and commercialization process.

References

Abadie, A. and G. Imbens (2006) "Large Sample Properties of Matching Estimators for Average Treatment Effects", *Econometrica*, Vol. 74, No. 1, pp. 235-267.

Angrist, J. and V. Lavy (2001) "Does Teacher Training Affect Pupil Learning? Evidence from Matched Comparisons in Jerusalem Public Schools", *Journal of Labor Economics*, Vol. 19, No. 2, pp. 343-369.

Brander, J., R. Amit, and W. Antweiler (2002) "Venture-Capital Sundication: Improved Venture Selection vs. the Value-Added Hypothesis", *Journal of Economics and Management Strategy*, Vol. 11, No. 3, pp. 423-452.

Brander, J., E. Egan, and T. Hellman (2010) "Government Sponsored versus Private Venture Capital: Canadian Evidence", in J. Lerner and A. Schoar, eds., *International Differences in Entrepreneurship*, NBER Conference Report, University of Chicago Press pp. 275-320.

Berube, C. and P. Mohnen (2009) "Are Firms that Receive R&D Subsidies More Innovative?", *Canadian Journal of Economics*, Vol. 42, No. 1, pp. 206-225.

Cochran, W. and D. B. Rubin (1973) "Controlling Bias in Observational Studies", *Sankyha: The Indian Journal of Statistics*,, Vol. 35, pp. 417-446.

Da Rin, M. and M. Penas (2007) "The Effect of Venture Capital on Innovation Strategies", NBER Working Paper no. w13636.

Engel, D. and M. Keilback (2007) "Firm Level Implication of Early Stage Venture Capital Investments: an Empirical Investigation", *Journal of Empirical Finance*, Vol 14, pp. 150-167.

Heckman, J., H. Ichimura, and P. Todd (1997) "Matching as an Econometric Evaluation Estimator: Evidence from Evaluating a Job Training Program", *Review of Economic Studies*, Vol. 64, No. 2, pp. 605-654.

Heckman, J., H. Ichimura, and P. Todd (1997) "Matching as an Econometric Evaluation Estimator", *Review of Economic Studies*, Vol. 65, No. 4, pp. 261-294.

Hellman, T. and M. Puri (2000) "The Interaction between Product Market and Financing Strategy: The Role of Venture Capital", *Review of Financial Studies*, Vol. 13, pp. 959-984.

Hellman, T. and M. Puri (2002) "Venture Capital and the Professionalization of Start-up Firms: Empirical Evidence", *Journal of Finance* Vol. 57, pp. 169-197.

Jain, B. A. and O. Kini (1995) "Venture Capitalist Participation and the Post-Issue Operating Performance of IPO Firms", *Managerial and Decision Economics*, Vol. 16, No. 6, pp. 596-606.

Kortum, S. and J. Lerner (2000) "Assessing the Contribution of Venture Capital and Innovation: Which is First?", *RAND Journal of Economics*, Vol. 31, No. 4, pp. 674-692.

Lavy, V. (2009) "Performance Pay and Teachers' Effort, Productivity, and Grading Ethics", *American Economic Review*, Vol. 99, No. 5, pp. 1979-2011.

Lerner, J. (2010) *Boulevard of Broken Dreams: Why Public Efforts to Boost Entrepreneurship and Venture Capital Have Failed - and What to Do about It*. Kauffman Foundation Series on Innovation and Entrepreneurship. Princeton University Press, Princeton, NJ.

Puri, M. and R. Zarutskie (2009) "On the Lifecycle Dynamics of Venture-Capital- and Non-Venture-Capital-Financed Firms", NBER Working Paper no. w14250.

Rosenbaum, P. and D. Rubin (1983) "The Central Role of the Propensity Score in Observational Studies for Causal Effects", *Biometrika,* Vol. 70, No. 1, pp. 41-55.

Rosenbaum, P. and D. Rubin (1985) "Constructing a control group using multivariate matched sampling incorporating the propensity score", *The American Statistician,* Vol. 39, pp. 33-38.

Rubin, D. B. (1976) "Multivariate matching methods that are equal percent bias reducing, I: some examples", *Biometrics,* Vol. 32, pp. 109-120.

Rubin, D. B. (1980) "Bias Reduction using Mahalanobis Metric Matching", *Biometrics,* Vol. 36, pp. 293-298.

Rubin, D. B. (2001) "Using Propensity Scores to Help Design Observational Studies: Application of the Tobacco Litigation", *Health Services & Outcomes Research Methodology,* Vol. 2, pp. 169-188.

Samila, S. and O. Sorenson (2011) "Venture Capital, Entrepreneurship, and Economic Growth", *Review of Economics and Statistics,* Vol. 93, pp. 338-349.

Wooldridge, J. M. (2010) *Econometric Analysis of Cross Section and Panel Data, Second Edition*. The MIT Press, Cambridge, MA.

Annex A: Data Linkage

This paper relies on new data linking information on VC activity in Canada provided by Thomson Reuters and information on IRAP recipients to Statistics Canada administrative databases. The initial linkage process matched these sets of firms to the BR based on their name and address, achieving varying degrees of quality in terms of the similarity of the firm's name and address between the external and STC data sets. The linkage rates and quality of these linkages for the VC data and IRAP data are reported in Table A1 and A3, respectively.

Table A1: Linkage of Thomson Reuters VC data to the BR

<table>
<tr><th>Name</th><th>Address</th><th>Postal Code</th><th>City</th><th>Quality Flag</th><th>Counts</th><th>Percentage</th><th>Cumulative Percentage</th></tr>
<tr><td rowspan="4">Same</td><td colspan="3">Same</td><td>D1</td><td>208</td><td>8%</td><td>8%</td></tr>
<tr><td>Same</td><td></td><td></td><td>D2</td><td>60</td><td>2%</td><td>10%</td></tr>
<tr><td></td><td>Same</td><td></td><td>D3</td><td>812</td><td>29%</td><td>39%</td></tr>
<tr><td></td><td></td><td>Same</td><td>D4</td><td>227</td><td>8%</td><td>47%</td></tr>
<tr><td rowspan="4">Similar</td><td colspan="3">Same</td><td>M1</td><td>32</td><td>1%</td><td>48%</td></tr>
<tr><td>Same</td><td></td><td></td><td>M2</td><td>69</td><td>2%</td><td>51%</td></tr>
<tr><td></td><td>Same</td><td></td><td>M3</td><td>96</td><td>3%</td><td>54%</td></tr>
<tr><td></td><td></td><td>Same</td><td>M4</td><td>86</td><td>3%</td><td>58%</td></tr>
<tr><td colspan="4">Manual Match</td><td></td><td>463</td><td>17%</td><td>74%</td></tr>
<tr><td colspan="4">Unmatched</td><td></td><td>709</td><td>26%</td><td>100%</td></tr>
<tr><td colspan="4">Total</td><td></td><td>2762</td><td>100%</td><td></td></tr>
</table>

Among the VC data, a number of the VC-backed firms had suspected linkages in the BR, but these potential counterparts were not identical in either their name or address. This is not altogether surprising as the data collection methods used by Thomson Reuters do not necessarily produce the legal name of the firm. Further, there is the potential for these VC-backed firms to alter their name during the lag from when the firm is captured in either dataset. For these records, analysts from STC and the Small Business Branch of Industry Canada made these linkages manually, leveraging their expertise and public sources of information on these firms.

For the VC data, following the initial match, information from other STC data sets including T2, T4 and PD7 remittance slips were used to further assess the quality of these initial matches. Following a guiding prinicipal of conservative accuracy, suspect linkages were identified and removed from the data based on a set of prudent rules. In particular, linked records were removed if: the VC-backed firm is not identified in the BR within two years of the first round of financing;[10] the firm reported no economic activity in the year of or year following investment;[11] the firm reported revenues exceeding $50 million at the time of first financing; and whether there were inconsistencies between the STC administrative data and the VC exit data supplied by Thomson Reuters. After applying these rules, we are left with a sample 1,545 VC-backed firms that were successfully linked to the STC data. Table A2 summarizes attrition from each of the rules described above applied sequentially.

[10] Firms receiving their first round of VC financing prior to 1999 were removed if they were not present in the BR in 1999 or 2000.
[11] Economic activity is defined as filing a T2, T4, or PD7 payroll remittance forms. Firms receiving their first round of VC financing prior to 1999 were removed if they showed no signs of economic activity in 1999 and 2000.

Table A2: Filtering Process and Attrition in Developing the Final Research Database and Treatment Group

	Restrictions	Attrition	Sample Size
Thompson Reuters List of VC-Backed Firms (1990 – 2009)			**2,762**
	Successful matches to the Statistics Canada Business Registry (BR)	709	2,053
	BR match in first year of financing or following year	372	1,681
	Economic activity (T2, T4, or PD7) in year of financing or following year	81	1,600
	Outliers at time of first financing	11	1,589
Research database consisting of 1545 VC-Backed firms (1999 – 2009)	Exit Data Analysis	44	**1,545**
	Received first round of VC financing during 1999-2009	501	1,044
	Reported sufficient financial information in in-scope industries	382	662
Treatment group of VC-backed firms	Successfully matched with comparable non VC-backed firm	118	**544**

The linkage process for the VC data was designed to accommodate various forms of research and analysis. While there are 1,545 VC-backed firms operating in Canada during 1999-2009, many of these firms received VC financing prior to the start of the sample, and as a result, are outside the scope of this study. The research database contains 1,044 firms that received their first round of VC financing during the 1999-2009 sample period.

Table A3: Linkage of IRAP data to the BR

Name	**City**	**Province**	**Quality Flag**	**Counts**	**Percentage**	**Cumulative Percentage**
Same	Same		D1	5187	67%	67%
	Same		D2	961	12%	79%
		Same	D3	6	0%	80%
Similar	Same		M1	192	2%	82%
		Same	M3	122	0%	84%
Unmatched				1268	15%	100%
Total				7736	100%	

The IRAP data contained less information on the operating address of the firm and as a result, we imposed stronger criteria for determining a linkage. However, the IRAP data had a much higher quality linkage with 80 per cent of the sample reporting an identical name in both the IRAP data and the BR. The relatively higher congruence among the company names in the IRAP and STC data could be related to their respective data collection processes. While firms directly supply the National Research Council with their legal names in applying for IRAP, the Thomson VC data is indirectly collected through the investing VC funds and Limited Partners. Given the high degree of concordance among the STC and IRAP data, we did not pursue further means to improve the linkage.

Annex B: Labour Tracking

As mentioned in section 2.2, the BRID assigned to enterprises in the BR is not designed to track a given enterprise over time. In particular, an enterprise's BRID can change for relatively benign reasons, e.g., a change in their legal name. Conversely, an enterprise could maintain the same BRID even following substantial structural changes including merger and acquisition activitiy.

To mitigate potential biases in subsequent growth estimates using these data, analysts at STC conducted a labour-tracking methodology in conjunction with IC analysts familiar with VC-backed companies. The procedure involved following masses of employment (i.e., individuals identified by their SINs in the T4 tax files) moving from one BRID to another. Depending on the nature of these relationships, we ammended the respective BRID entries to reflect a single longitudinal record for each firm within the treatment and control groups using the year they were matched as a base year.

Under labour tracking, relationships among the BRIDs are only identified when one BRID either starts or stops filing T4 tax information—a T4 birth or T4 death, respectively. Once a T4 death/birth occurs, whether there is sufficient evidence of a relationship with another BRID depends on the size of the firm and the proportion of shared employees among the predecessor and successor BRIDs. The thresholds to determine whether a relationship exists are summarized in Table B1. Approximately one third of the BRIDs associated with firms in the treatment and control groups were involved in one or more labour tracking relationships meeting these criteria.

Table B1: Thresholds for Shared Employment to Identify Labour Tracking Relationships

	Size of T4 Birth / Death					
	> 250 Employees	**> 50 Employees**	**> 15 Employees**	**> 7 Employees**	**> 5 Employees**	**5 Employees**
Proportion of Shared Employees	25%	30%	50%	50% if target is a birth / death 60% if target is a continuer	70%	100%

There are three basic types of labour-tracking relationships: a death-to-birth; a death-to-continuer; and a continuer-to-birth. While specific events in the firm life cycle cannot be robustly inferred from this methodology, these relationships roughly translate into a false death (i.e., same firm reports under a new BRID), an acquisition, and a divestiture or spin-off. Further, a BRID may be involved in multiple relationships, e.g., two deaths to a birth roughly corresponds to a merger. These relationships can become quite complex when a BRID is involved with many different types of relationships spanning several other BRIDs, a situation which is not altogether uncommon among BRIDs corresponding to larger firms.

For our purpose, we do not want to focus solely on 'organic' growth within firms.[12] Our objective is to measure the performance of VC-backed firms, and growth through acquisition could be part of their strategy to expand their operations. At the same time, if the VC-backed firm is the target of a merger or acquisition, this is likely an exit for the VC investor, and ideally the end of the longitudinal record. Regarding spin-offs, it remains unclear whether the new BRID reflects a separate entity or a change in the firm's reporting practices. Accordingly, we connect records when labour tracking suggests there has been a false death, and end records when labour tracking suggests the treatment/control firm is acquired or that there was a substantial spin-off.

We apply a 50-per-cent shared employee threshold for evaluating all labour tracking relationships.[13] In particular, for a potential false death, the predecessor and successor BRIDs are connected into one longitudinal record if the successor BRID shares 50 per cent or more of the employees of the predecessor. For cases suggesting the treatment/control firm acquired another firm, we end the record if the treatment/control firm represents less than 50 per cent of the shared employment in the merged firm. For potential divestitures, we end the record if the spin-off exceeds 50 per cent of the employment of the treatment/control firm. Finally, for complex cases where a BRID is involved in many relationships, we apply the 50-per-cent threshold cumulatively, ending the record if the total employment associated with all acquisitions and divestitures exceeds 50 per cent of the treatment/control firm's employment.

The Thomson VC data also contains information on VC exits including IPOs, M&A activity, and business failures. These data are used to supplement the exits identified through labour tracking among the treatment group.

[12] 'Organic' growth in reference to employment generally refers to net job creation or destruction independent of employment reallocation. However, among our data, employment changes within the firm include non 'organic' changes, e.g., mergers, acquisitions, and divestitures. While non 'organic' growth may amount to nothing more than job shuffling when viewed at the industry level, these aspects of growth reflect changes in capacity at the firm level.

[13] Although somewhat *ad-hoc*, this threshold was selected in consultation with Statistics Canada and Industry Canada analysts to indicate that the treatment/control firm is the dominant component of the labour tracking relationship. Thresholds beyond 50 per cent were found to be too strict given regular rates of employee attrition and reallocation across firms.

Annex C: Estimating the Predicted Probability of Receiving VC

As explained in section 3, we estimate the predicted probability using several firm characteristics that could affect VC financing decisions, and the squared values of these covariates to allow for non-linear relationships. The estimation results are reported in Table C1. We calculate the propensity scores for our matching results through these logit regression estimates.

The coefficients for the measures on firm size, i.e., total assets, employment, sales and revenue, suggest that firm size increases the probability that a given firm will receive VC financing, but only until they reach a certain threshold as the coefficients for their squared terms are all negative. Interestingly, sales and revenue seem to impact the propensity score less than the other scale measures and are also less statistically significant. This could result from a high degree of collinearity among the size measures, or could imply that VC funds invest at various stages of the firm life cycle, some of which is associated with low sales.

As one might expect given the ratios shown in Table 1, the results also suggest that VC funds typically invest in younger firms. Further, when controling for firm size, net income has a negative impact on the probability of receiving financing. Similarly, retained earnings also appear to exert a negative influence on the propensity score. The negative results for measures related to profitability likely stem from VC's penchant for early-stage firms, as well as a selection effect where firms with large profits can self-finance and/or secure financing better through more traditional channels.

Turning to our measures on firm innovation, we find that they are highly significant and contribute substantially to the probability of receiving VC. Becoming an IRAP recipient increases the propensity score for a given firm by over 300 per cent. Although marginal changes in R&D expenditures did not have a particularly large impact on the probability of receiving VC, the discreet change from a non-R&D performer to an R&D performer greatly improved the chances for the treatment. For example, among the sample of the non-R&D performers in the treatment and control groups, a discreet change in expenditures from 0 to $15,000 (a relatively small amount) led to a 460-per-cent increase in their probabiltiy of receiving VC.

Table C1: Logit - Pr(VC first financing = 1)

Independent Vars	Coeff.	Std Error	P-value
Ln Total Assets	3.527	0.384	0.000
H Sales	0.021	0.038	0.583
Ln Employment	0.538	0.107	0.000
Ln Wages	6.270	1.736	0.000
H Retained Earnings	-0.043	0.005	0.000
H Revenue	0.018	0.014	0.200
H Net Income	-0.029	0.005	0.000
Age	-0.076	0.012	0.000
H R&D Expenditures	0.270	0.045	0.000
Ln Total Assets Sq	-0.107	0.013	0.000
H Sales sq	-0.002	0.003	0.415
Ln Employment sq	-0.056	0.019	0.003
Ln Wages sq	-0.282	0.082	0.001
H Retained Earnings Sq	-0.008	0.001	0.000
H Revenue sq	-0.009	0.002	0.000
H Net Income sq	0.014	0.002	0.000
Age sq	0.001	0.000	0.021
H R&D Expenditures Sq	-0.009	0.003	0.004
IRAP dummy	1.154	0.097	0.000
R&D 2000 dummy	1.973	0.253	0.000
Industry effects	Yes		
Year effects	Yes		
N	2,573,663	Pr>ChiSq	0.000
LR test	5059.62	LL	-3603.92
Pseudo R2	0.412		

The sample comprises all firms filing a T2 and T4 spanning 1999-2009 within NAICS 4 digit industries where at least three firms have received VC financing at some point during the 1990-2009 period.

All values are in either natural logs (Ln), or the inverse hyperbolic sine transformation (H).

Annex D: Robustness Check

Table D1: Differences in Growth Rates Among Matched Pairs of the Treatment and Control Group where both Report Financials over 1, 3 and 5 years

	Mean VC	Mean Control Group	Difference	P-value of Difference	N VC	N Control Group
Total Assets						
1 Yr growth	0.208	0.016	0.192	0.001	455	455
3 Yr growth	0.369	-0.101	0.470	0.000	255	255
5 Yr growth	0.472	0.035	0.437	0.046	133	133
Total Revenue						
1 Yr growth	0.490	0.302	0.189	0.017	415	415
3 Yr growth	0.858	0.418	0.440	0.005	226	226
5 Yr growth	0.967	0.566	0.401	0.149	114	114
Sales						
1 Yr growth	0.478	0.285	0.192	0.062	304	304
3 Yr growth	0.961	0.470	0.492	0.014	159	159
5 Yr growth	1.412	0.589	0.823	0.015	72	72
Employment						
1 Yr growth	0.346	0.067	0.278	0.000	458	458
3 Yr growth	0.436	-0.058	0.494	0.000	233	233
5 Yr growth	0.474	0.052	0.422	0.008	106	106
Wages						
1 Yr growth	0.084	0.053	0.031	0.120	458	458
3 Yr growth	0.165	0.112	0.053	0.155	233	233
5 Yr growth	0.339	0.143	0.196	0.006	106	106
R&D Expenditures						
1 Yr growth	0.245	0.065	0.179	0.009	234	234
3 Yr growth	0.176	-0.007	0.182	0.334	101	101
5 Yr growth	0.367	0.079	0.288	0.348	37	37
Gross margin						
1 Yr growth	-0.005	-0.009	0.004	0.864	236	236
3 Yr growth	0.018	0.112	-0.094	0.237	129	129
5 Yr growth	-0.001	0.049	-0.050	0.437	58	58
Gross Profit/Employment						
1 Yr growth	0.187	0.276	-0.088	0.409	231	231
3 Yr growth	0.638	0.411	0.227	0.213	120	120
5 Yr growth	0.657	0.794	-0.136	0.660	55	55